

CONVEX HULLS OF SUPERINCREASING KNAPSACKS AND LEXICOGRAPHIC ORDERINGS

AKSHAY GUPTE

Department of Mathematical Sciences, Clemson University

ABSTRACT. We consider bounded integer knapsacks where the weights and variable upper bounds together form a superincreasing sequence. The elements of this superincreasing knapsack are exactly those vectors that are lexicographically smaller than the greedy solution to optimizing over this knapsack. We describe the convex hull of this n -dimensional set with $\mathcal{O}(n)$ facets. We also establish a distributive property by proving that the convex hull of \leq - and \geq -type superincreasing knapsacks can be obtained by intersecting the convex hulls of \leq - and \geq -sets taken individually. Our proofs generalize existing results for the $0\backslash 1$ case.

1. INTRODUCTION

Given positive integers n, b , and (a_i, u_i) for all $i \in N := \{1, \dots, n\}$, we consider a bounded integer knapsack defined as $K := \{x \in \mathbb{Z}_+^n : a^\top x \leq b, 0 \leq x_i \leq u_i \text{ } i = 1, \dots, n\}$. Without loss of generality (w.o.l.o.g.) we assume that $a_i u_i \leq b \forall i \in N$ and that $a^\top u > b$ to ensure a nontrivial set. When all the upper bounds are equal to one, we have the $0\backslash 1$ knapsack K^1 . The convex hull of K , denoted by $\text{conv } K$, is referred to as the knapsack polytope.

The study of the knapsack polytope has received considerable attention in literature and in general, there may exist exponentially many facet-defining inequalities. There exist special classes of K for which a complete description of $\text{conv } K$ is known. For the $0\backslash 1$ knapsack, these results include the minimal covers of Wolsey [1] assuming certain matroidal properties for $a^\top x \leq b$, $(1, k)$ -configurations of Padberg [2], weight-reduction principle of Weismantel [3] when $a_i \in \{1, \lfloor b/3 \rfloor + 1, \lfloor b/3 \rfloor + 2, \dots, \lfloor b/2 \rfloor\} \forall i$ or $a_i \in \{1, \lfloor b/2 \rfloor + 1, \lfloor b/2 \rfloor + 2, \dots, b\} \forall i$, and Weismantel [4] when $a_i \in \{\acute{a}, \check{a}\} \forall i$ and for two distinct positive integers \acute{a} and \check{a} . There also exist complete descriptions of $\text{conv } K$ for upper bounds not equal to 1. For a *divisible* knapsack, i.e. when $a_{i-1} \mid a_i$ for all $i \geq 2$, three results are known: (i) Marcotte [5] when $K = \{x \in \mathbb{Z}_+^n : a^\top x \leq b\}$, (ii) Pochet and Wolsey [6] when $K = \{x \in \mathbb{Z}_+^n : a^\top x \geq b\}$, and (iii) Pochet and Weismantel [7] when $K = \{x \in \mathbb{Z}_+^n : a^\top x \leq b, \mathbf{0} \leq x \leq u\}$. Recently, Cacchiani et al. [8] described the convex hulls of \leq - and \geq -type knapsacks with a generalized upper bound constraint $\sum_{i \in N} x_i \leq 2$. The polytopes in [6, 7] involve an exponential number of valid inequalities, whereas the polytopes in [5, 8] have $\mathcal{O}(n)$ facets.

In this paper, we are interested in the convex hull of a special class of K characterized as follows.

Definition 1 (Superincreasing knapsack). The set K is said to be a *superincreasing knapsack* if $\{(a_i, u_i)\}_{i \in N}$ forms a weakly superincreasing sequence of tuples, i.e. $\sum_{k=1}^i a_k u_k \leq a_{i+1} \forall i \geq 1$.

E-mail address: agupte@clemson.edu.

Key words and phrases. superincreasing sequence, lexicographic ordering, greedy solution knapsack, convex hull, linear time complexity.

0\1 superincreasing knapsacks have been historically used in cryptographic systems [9, 10]; their structure and linear time complexity were investigated in [11]. Atkinson et al. [12] obtained a wider class of 0\1 knapsacks with linear time complexity. From the viewpoint of a polyhedral study, the knapsack polytope for superincreasing K^1 was first studied by Laurent and Sassano [13], whose result is paraphrased below.

Theorem 1 ([13]). *For any positive integer b , the 0\1 knapsack polytope $\text{conv } K^1[n][b]$ is completely described by its minimal cover inequalities if and only if $\{(a_i, 1)\}_{i \in N}$ forms a weakly superincreasing sequence. Furthermore, all the $\mathcal{O}(n)$ minimal covers can be explicitly enumerated.*

We extend the sufficiency condition of Theorem 1 to the case when the variable upper bounds in K are not necessarily equal to one. We explicitly describe the convex hull with $\mathcal{O}(n)$ nontrivial facets. As is to be expected, the proposed inequalities reduce to minimal covers of K^1 when u is a vector of ones. The convex hull proof for K^1 is simpler since the coefficient matrix for system of minimal covers is an interval matrix and hence totally unimodular. Since Marcotte [5] describes the convex hull of divisible K using $\mathcal{O}(n)$ facets, it is obvious that the superincreasing property is not a necessary condition for $\text{conv } K$ to have $\mathcal{O}(n)$ facets. Besides generalizing the result of Laurent and Sassano, another motivation for studying superincreasing knapsacks is that such sets appear after reformulating the integer variables in a mixed integer program; see Gupte et al. [14]. The most common example of such reformulations is the set

$$K_\alpha(b) := \left\{ \zeta : \sum_t \alpha^{t-1} \zeta_t \leq u, \zeta_t \in \{0, 1, \dots, \alpha - 1\} \quad \forall t = 1, \dots, \lfloor \log_\alpha u \rfloor + 1 \right\} \quad (1)$$

obtained after α -nary expansion of a integer variable: $x = \sum_{t=1}^{\lfloor \log_\alpha u \rfloor + 1} \alpha^{t-1} \zeta_t$. Convex hull of $K_2(b)$ was independently studied by [14, 15]. A complete knowledge of the superincreasing knapsack polytope will provide a family of valid inequalities to the mixed integer program. Gupte et al. demonstrated the practical usefulness of facets to binary expansion knapsacks as cutting planes in a branch-and-cut algorithm for solving mixed integer bilinear programs.

Remark 1. The extended formulation of K , obtained after adding new variables $\zeta_{it} \in \{0, 1\} \forall i, t$ and basis expansion of each x_i as $x_i = \sum_{t=1}^{\lfloor \log_\alpha u_i \rfloor + 1} \alpha^{t-1} \zeta_{it}$ for some $\alpha \in \mathbb{Z}_{++}$, does not obey the superincreasing property. Hence we cannot obtain $\text{conv } K$ simply as a projection of the extended formulation.

Note that there is no inclusive relationship between superincreasing and divisible knapsacks. However, certain types of knapsacks, such as $K_\alpha(b)$, may be both divisible and superincreasing. For divisible superincreasing knapsacks, our result provides a explicit linear size minimal description as compared to the implicit exponential size description in [7].

Throughout this paper, we assume that K is superincreasing. We begin by analyzing the greedy solution of K in Section 2 and use it to provide a useful geometric interpretation to our assumption of superincreasing tuples $\{(a_i, u_i)\}_{i \in N}$. Section 3 derives a set of facet-defining inequalities, referred to as *packing inequalities*, to $\text{conv } K$ and our first main result in Theorem 2 proves that these inequalities describe $\text{conv } K$. In Section 4, we prove that the convex hull of intersection of two superincreasing knapsacks is given by the facets of the individual knapsack polytopes. This second main result in Theorem 3 is indeed interesting since the convex hull operator does not distribute in general and further implies that the convex hull of a family of m intersecting superincreasing knapsacks is described by $\mathcal{O}(n)$ linear inequalities. For general 0\1 knapsacks, new valid inequalities were derived in [16, 17] for intersection of two \leq -type knapsacks and in [17, 18] for one \leq - and one \geq -type knapsack.

We adopt the following notation. $\text{conv } \mathcal{X}$ is the convex hull of a set \mathcal{X} . $\mathbb{Z}_+(\mathbb{Z}_{++})$ is the set of nonnegative (positive) integers. \mathbf{e} is a vector of ones, \mathbf{e}_i is the i^{th} unit vector and $\mathbf{0}$ is a vector of zeros. $\mathcal{H}_n(u) := \{x \in$

$\mathbb{Z}_+^n: x_i \leq u_i \forall i$ is a discrete hyper-rectangle. For $\ell > s$, we denote $\sum_\ell^s(\cdot) = 0$ and $\prod_\ell^s(\cdot) = 1$. The positive part of $\xi \in \mathbb{R}$ is denoted by $[\xi]^+ := \max\{0, \xi\}$.

2. STRUCTURE OF K

This section discusses structural properties of a superincreasing knapsack - first we present a geometric interpretation to the algebraic requirements of Definition 1, then we characterize maximal packings of K and finally we state a dynamic program to optimize over K . The proposed results, especially the maximal packing and dynamic program, are known in literature for the $0/1$ case, see for example Shamir [11], which has important applications in cryptographic systems. Our contribution is to extend these results to the general integer case and establish a foundation for our main theorems in Sections 3 and 4.

The notion of *lexicographic ordering* will be useful for the rest of the paper. For any two vectors v^1 and v^2 , the vector v^1 is lexicographically smaller than v^2 , denoted as $v^1 \preceq v^2$, if either $v^1 = v^2$ or the first (in reverse order) nonzero element i of $v^1 - v^2$ is such that $v_i^1 < v_i^2$. In the latter case, we denote $v^1 \prec v^2$. Since \preceq is a total order, for any distinct v^1 and v^2 , either $v^1 \prec v^2$ or $v^2 \prec v^1$ (equivalently $v^1 \succ v^2$).

The *greedy solution* of an arbitrary (not necessarily superincreasing) knapsack $K^{\text{gen}} := \{x \in \mathcal{H}_n(\tilde{u}) : \tilde{a}^\top x \leq \tilde{b}\}$ is given by

$$\tilde{\theta}_i := \min \left\{ \tilde{u}_i, \left\lfloor \frac{\tilde{b} - \sum_{k=i+1}^n \tilde{a}_k \tilde{\theta}_k}{\tilde{a}_i} \right\rfloor \right\} \quad \forall i = n, \dots, 1. \quad (2)$$

Magazine et al. [19] referred to this solution in the case of trivial upper bounds, i.e., when $\tilde{u}_i = \lfloor \tilde{b}/\tilde{a}_i \rfloor \forall i$, and studied conditions under which it is the optimal solution for maximizing over K^{gen} . By construction, we have $\tilde{\theta} \in K^{\text{gen}}$. In fact, $\tilde{\theta}$ is lexicographically the largest vector in K^{gen} .

Lemma 1. $K^{\text{gen}} \subseteq \{x \in \mathcal{H}_n(\tilde{u}) : x \preceq \tilde{\theta}\}$.

Proof. Suppose there exists some $x \in K^{\text{gen}}$ with $x \succ \tilde{\theta}$. Let $i^* := \max\{i \in N : x_i \neq \tilde{\theta}_i\}$. Since $x \succ \tilde{\theta}$ by assumption and $x, \tilde{\theta} \in \mathbb{Z}^n$, we have $x_{i^*} \geq \tilde{\theta}_{i^*} + 1$. Now $x_{i^*} \leq \tilde{u}_{i^*}$ and equation (2) imply that $\tilde{\theta}_{i^*} = \lfloor \frac{\tilde{b} - \sum_{i>i^*} \tilde{a}_i \tilde{\theta}_i}{\tilde{a}_{i^*}} \rfloor$. Hence $x_{i^*} > (\tilde{b} - \sum_{i>i^*} \tilde{a}_i \tilde{\theta}_i) / \tilde{a}_{i^*}$. Then $\tilde{a}^\top x \geq \sum_{i>i^*} \tilde{a}_i x_i + \tilde{a}_{i^*} x_{i^*} > \sum_{i>i^*} \tilde{a}_i \tilde{\theta}_i + \tilde{b} - \sum_{i>i^*} \tilde{a}_i \tilde{\theta}_i = \tilde{b}$, a contradiction to $x \in K^{\text{gen}}$. \square

Henceforth, we denote θ to be the greedy solution of a superincreasing knapsack K . Proposition 1 states that for superincreasing knapsacks, the inclusion in Lemma 1 becomes an equality. The proof of this depends on the following observation about points in K .

Lemma 2. Let $x \in K$ and $y \in \mathcal{H}_n(u)$ be such that $y \preceq x$. Then $a^\top y \leq a^\top x$ and hence $y \in K$. If $y \prec x$ and $i^* := \max\{i \in N : y_i \neq x_i\}$, we have

- (1) $a^\top y = a^\top x$ if and only if $a_{i^*} = \sum_{i<i^*} a_i u_i$, $y_{i^*} = x_{i^*} - 1$, and $y_i = u_i, x_i = 0 \forall i < i^*$.
- (2) $(u_1, \dots, u_{i^*-1}, y_{i^*}, x_{i^*+1}, \dots, x_n) \in K$.

Proof. Suppose that $y \neq x$. Since $y \preceq x$ and $y \in \mathbb{Z}_+^n$, we have $y_{i^*} \leq x_{i^*} - 1$. Then, $a^\top(y - x) = \sum_{i<i^*} a_i(y_i - x_i) + a_{i^*}(y_{i^*} - x_{i^*}) \leq \sum_{i<i^*} a_i u_i - a_{i^*} \leq 0$, giving us $a^\top y \leq a^\top x$ and $y \in K$. This also leads to the conditions for $a^\top(y - x) = 0$. Finally, $(u_1, \dots, u_{i^*-1}, y_{i^*}, x_{i^*+1}, \dots, x_n) \prec x$ implies the second claim. \square

Lemmas 1 and 2 and the fact that $\theta \in K$ gives us

Proposition 1. $K = \{x \in \mathcal{H}_n(u) : x \preceq \theta\}$.

Thus, a superincreasing knapsack is exactly the set of integer points within $\mathcal{H}_n(u)$ that are lexicographically smaller than the greedy solution. This implies that two distinct superincreasing sequences, $\{a, u\}$ and $\{w, u\}$, represent the same knapsack if and only if the corresponding greedy solutions are equal.

We now describe the connection between the greedy solution and the notion of *maximal packing*. Let $g(\tilde{a}, \tilde{u}, \tilde{b}) := \max\{\tilde{a}^\top x : \tilde{a}^\top x \leq \tilde{b}, x \in \mathcal{H}_n(\tilde{u})\}$ denote the maximum attainable capacity of K^{gen} . Clearly, $K^{\text{gen}} = \{x \in \mathcal{H}_n(\tilde{u}) : \tilde{a}^\top x \leq g(\tilde{a}, \tilde{u}, \tilde{b})\}$. A maximal packing of K^{gen} is a vector $x \in K^{\text{gen}}$ such that $\tilde{a}^\top x = g(\tilde{a}, \tilde{u}, \tilde{b})$. Maximal packing may not be unique and in general, computing it reduces to solving the NP-hard subset-sum problem, although it can be computed in linear time for divisible knapsacks [20]. For superincreasing knapsacks, Proposition 1 and Lemma 2 imply that θ is a maximal packing of K . The conditions for θ being the only maximal packing are characterized next.

Proposition 2. *θ is a unique maximal packing of K if and only if for every $j \in N$ with $a_j = \sum_{i=1}^{j-1} a_i u_i$ and $\theta_j > 0$, there exists $i < j$ such that $\theta_i > 0$.*

Proof. Assume that for every $j \in N$ with $a_j = \sum_{i=1}^{j-1} a_i u_i$, there exists $i < j$ such that $\theta_i > 0$. Suppose θ is not the unique maximal packing and there exists some $\gamma \in K \setminus \{\theta\}$ such that $a^\top \gamma = a^\top \theta$. Proposition 1 gives us $\gamma \prec \theta$. Then Lemma 2 implies that $a_{i^*} = \sum_{i < i^*} a_i u_i$ and $\theta_i = 0 \forall i < i^*$, a contradiction to our assumption. Now suppose that θ is unique and let there exist some $j \in N$ with $a_j = \sum_{i=1}^{j-1} a_i u_i$ and $\theta_j > 0$ but $\theta_i = 0$ for all $i < j$. Set $\gamma \prec \theta$ as follows: $\gamma_i = u_i \forall i < j$, $\gamma_j = \theta_j - 1$, and $\gamma_i = \theta_i \forall i > j$. Then $\gamma \in \mathcal{H}_n(u)$ and $a^\top \gamma = a^\top \theta = g(a, u, b)$, contradicting the uniqueness of θ . \square

Remark 2. When $a_i = \alpha^{i-1}$ and $u_i = \alpha - 1$ for all $i \in N$ and some $\alpha \in \mathbb{Z}_{++}$, we have the set $K_\alpha(b)$ from (1) and it is straightforward to verify in this case that θ is the unique representation of b in base α .

Example 1. Let $a = (2, 8, 46, 150, 310)$, $u = (3, 5, 2, 1, 2)$ and $K = \{x \in \mathcal{H}_5(u) : \sum_{i=1}^5 a_i x_i \leq 841\}$. It can be verified by enumerating the points in K that $g(a, u, 841) = 840$. The greedy solution is $\theta = (0, 3, 1, 1, 2)$ and observe that $a^\top \theta = 8(3) + 46 + 150 + 310(2) = 840$. Also, θ is the only point in K that yields $g(a, u, 841)$ and it satisfies the sufficient condition for uniqueness: $a_3 = a_1 u_1 + a_2 u_2$ with $\theta_3, \theta_2 > 0$. An alternate superincreasing knapsack representation is $K = \{x \in \mathcal{H}_5(u) : x_1 + 3x_2 + 18x_3 + 95x_4 + 189x_5 \leq 500\}$.

Now let $K' = \{x \in \mathcal{H}_5(u) : \sum_{i=1}^5 a_i x_i \leq 863\}$. We have $g(a, u, 863) = 862$, $\theta' = (0, 0, 2, 1, 2)$, $a^\top \theta' = 862$ and θ' does not satisfy the necessary condition for uniqueness: $a_3 = a_1 u_1 + a_2 u_2$ with $\theta'_3 > 0, \theta'_1 = \theta'_2 = 0$. Another maximal packing is $\gamma' = (3, 5, 1, 1, 2)$, which is equal to $\theta' + (u_1, u_2, -1, 0, 0)$.

Finally, let $\tilde{a} = (2, 8, 40, 150, 310)$, $\tilde{u} = (1, 5, 4, 1, 2)$ and note that $\tilde{a}_3 < \tilde{a}_1 \tilde{u}_1 + \tilde{a}_2 \tilde{u}_2$. For $K^{\text{gen}} = \{x \in \mathcal{H}_5(\tilde{u}) : \sum_{i=1}^5 \tilde{a}_i x_i \leq 825\}$ the greedy solution is $\tilde{\theta} = (1, 1, 1, 1, 2)$ with $\tilde{a}^\top \tilde{\theta} = 820 < 825 = g(\tilde{a}, \tilde{u}, 825) = \tilde{a}^\top \tilde{\gamma}$, where $\tilde{\gamma} = (1, 5, 4, 0, 2)$. Hence the greedy solution does not give a maximal packing. \diamond

Since θ is a maximal packing of K and is computable in $\mathcal{O}(n)$ time, we assume w.o.l.o.g. that $b = g(a, u, b) = a^\top \theta$.

The equivalence to lexicographic ordering in Proposition 1 lends intuition to the points contained in K and also enables us to prove our results. An immediate consequence is a linear time algorithm for optimization over K . To state this result, we first define the support of θ .

Definition 2. Let $I := \{i \in N : \theta_i \geq 1\}$ be the support of θ and denote¹ $I = \{i_1, \dots, i_r, i_{r+1} := n\}$ for some integer $r \geq 0$, where we assume $i_1 < i_2 < \dots < i_r < n$. For every $j \in N$, let $I_j := \{i \in I : i > j\}$, $I_j^- := \{i \in I : i < j\}$, $\text{prev}(j) := \max\{i : i \in I_j^-\}$, and $\text{next}(j) := \min\{i : i \in I_j\}$. If $j < i_1$ (resp. $j = n$), then $\text{prev}(j) = 0$ (resp. $\text{next}(j) = 0$).

¹ $n \in I$ since $a_n u_n \leq b$ implies $\theta_n = u_n$ and hence n is the largest index in I .

Clearly $I_n = I_{i_1}^- = \emptyset$ and $I_{j+1} = I_j \setminus \{j+1\}$ for all $j \in N$.

Proposition 3. *There exists a $\mathcal{O}(n)$ time algorithm to optimize over K . Given any $c \in \mathbb{R}^n$, for every $j \in I$, the optimal value $f_j^*(c) := \max\{\sum_{i=1}^j c_i x_i : x \in K, x_k = \theta_k \ \forall k \in I_j\}$ is equal to*

$$f_j^*(c) = \max \left\{ [c_j]^+(\theta_j - 1) + \sum_{i=1}^{j-1} [c_i]^+ u_i, c_j \theta_j + f_{\text{prev}(j)}^*(c) \right\}. \quad \square$$

The correctness of this recursion follows from Lemma 2, Proposition 1 and the observation that

$$\{x \in K : x_k = \theta_k \ \forall k \in I_j\} \subseteq \{x : x_j \in \{0, 1, \dots, \theta_j\}, x_i = 0 \ \forall i \in \{j+1, \dots, n\} \setminus I_j\}, \quad \forall j \in N. \quad (3)$$

Define $\Omega(c) := \arg \max\{c^\top x : x \in K\}$ as the set of optimal solutions to the maximization over K . The dynamic program of Proposition 3 can be represented as a binary tree with $|I| + 1$ leaf nodes, as illustrated in Figure 1. The elements of $\Omega(c)$ correspond to some of the leaf nodes whereas the elements of I are in a bijection to the non-leaf nodes. For a non-leaf $j \in I$, the set $\{x \in K : x_k = \theta_k \ \forall k \in I_j, x_j \leq \theta_j - 1\}$ contains the leaf descendant denoted by \mathcal{L}_j whereas the set $\{x \in K : x_k = \theta_k \ \forall k \in I_j, x_j = \theta_j\}$ corresponds to the next non-leaf node $\text{prev}(j)$ if $j \neq i_1$. The two leaf descendants of i_1 are \mathcal{L}_{i_1} and \mathcal{L}_0 . In particular,

$$\mathcal{L}_j = \{x : x_i \in \{0, u_i\} \ \forall i < j, x_j \in \{0, \theta_j - 1\}, x_i = \theta_i \ \forall i > j\} \quad \forall j \in I, \quad \text{and} \quad \mathcal{L}_0 = \{\theta\}. \quad (4)$$

This implies that

$$\Omega(c) \subseteq \{\theta\} \cup \bigcup_{j \in I} \mathcal{L}_j. \quad (5)$$

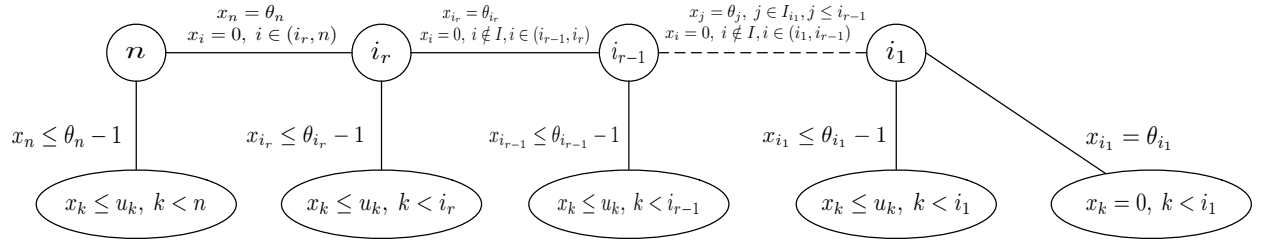


FIGURE 1. Binary tree for the dynamic programming algorithm of Proposition 3.

3. FACETS AND CONVEX HULL

We first derive valid inequalities whose coefficients depend on the greedy solution (and maximal packing) θ . To state the proposed *packing inequalities*, for any $j \in N \setminus n$, define a function $\phi_j : I_j \mapsto \mathbb{Z}_+$ as

$$\phi_j(i) := (u_j - \theta_j) \prod_{\substack{k=\text{next}(j): \\ k \in I}}^{\text{prev}(i)} (u_k + 1 - \theta_k) \quad i \in I_j. \quad (6)$$

From notational convention, we have $\phi_j(\text{next}(j)) = (u_j - \theta_j)$ and $\phi_n(\cdot) = 0$. The recursive definition of $\phi_j(\cdot)$ leads to the following identities that will be useful while arguing validity and facet-defining property.

Observation 1. *For any $j \in N \setminus n$ and $i \in I_j$, we have $\phi_j(\text{next}(i)) - \phi_j(i) = \phi_j(i)(u_i - \theta_i)$. Consequently, $\phi_j(i) = u_j - \theta_j + \sum_{k \in I_j : k < i} \phi_j(k)(u_k - \theta_k)$.*

Proof. The first statement is obvious from the definition of $\phi_j(\cdot)$. The second statement is obtained via a straightforward induction on i and using the first statement. \square

Proposition 4 (Packing inequalities). *For any $j \in N$, the inequality*

$$x_j + \sum_{i \in I_j} \phi_j(i)(x_i - \theta_i) \leq \theta_j \quad (7)$$

is valid to conv K .

Proof. For $j = n$, the inequality is simply $x_n \leq u_n$ since $I_n = \emptyset, u_n = \theta_n$. We prove validity for $j < n$ by induction on elements of I_j . Denote $\mathcal{S}_j := \{x \in K : x_k = \theta_k \ \forall k \in I_j\}$ for every $j \in N$. First we argue that $x_j + (u_j - \theta_j)(x_{\text{next}(j)} - \theta_{\text{next}(j)}) \leq \theta_j$ is valid for $\mathcal{S}_{\text{next}(j)}$. For $x \in \mathcal{S}_{\text{next}(j)}$, equation (3) gives us $x_{\text{next}(j)} \leq \theta_{\text{next}(j)}$. If $x_{\text{next}(j)} = \theta_{\text{next}(j)}$, then $x \in \mathcal{S}_j$ and the inequality reduces to $x_j \leq \theta_j$, which holds true by applying equation 3 to j . Otherwise, $x_{\text{next}(j)} \leq \theta_{\text{next}(j)} - 1$ and then $x_j + (u_j - \theta_j)(x_{\text{next}(j)} - \theta_{\text{next}(j)}) \leq x_j - u_j + \theta_j \leq \theta_j$.

Now we show that for any $k \in I_j$, $x_j + \sum_{i \in I_j : i \leq k} \phi_j(i)(x_i - \theta_i) \leq \theta_j$ is valid for \mathcal{S}_k . From the previous claim, the result is true for $k = \text{next}(j)$. Assume it to be true for some $k \in I_j$ and consider the inequality for $\mathcal{S}_{\text{next}(k)}$. For $x \in \mathcal{S}_{\text{next}(k)}$, equation 3 gives us $x_{\text{next}(k)} \leq \theta_{\text{next}(k)}$. If $x_{\text{next}(k)} = \theta_{\text{next}(k)}$, then $x \in \mathcal{S}_k$ and the inequality reduces to $x_j + \sum_{i \in I_j : i \leq k} \phi_j(i)(x_i - \theta_i) \leq \theta_j$, which is valid for \mathcal{S}_k from induction hypothesis. Otherwise, $x_{\text{next}(k)} \leq \theta_{\text{next}(k)} - 1$ and then

$$x_j - \theta_j + \sum_{i \in I_j : i \leq k} \phi_j(i)(x_i - \theta_i) + \phi_j(\text{next}(k))(x_{\text{next}(k)} - \theta_{\text{next}(k)}) \leq u_j - \theta_j + \sum_{i \in I_j : i \leq k} \phi_j(i)(u_i - \theta_i) - \phi_j(\text{next}(k)) = 0,$$

where the inequality is due to $\phi_j(\cdot) \geq 0$ and $x \leq u$ and the equality follows from Observation 1. This completes the induction process and our proof. \square

Since $\theta_j = u_j$ implies $\phi_j(i) = 0 \ \forall i \in I_j$, it follows that (7) reduces to $x_j \leq u_j$ when $\theta_j = u_j$. Thus the only nontrivial packing inequalities are those corresponding to $\theta_j < u_j$.

Example 1 (continued). Recall $K = \{x \in \mathbb{Z}_+^5 : 2x_1 + 8x_2 + 46x_3 + 150x_4 + 310x_5 \leq 841, x \leq (3, 5, 2, 1, 2)\}$ with $\theta = (0, 3, 1, 1, 2)$. We have $\phi_1(2) = 3, \phi_1(3) = 9, \phi_1(4) = \phi_1(5) = 18, \phi_2(3) = 2, \phi_2(4) = \phi_2(5) = 4, \phi_3(4) = \phi_3(5) = 1, \phi_4(5) = 0$. For $j \in \{1, 2, 3\}$, our maximal pack inequalities are

$$x_1 + 3x_2 + 9x_3 + 18x_4 + 18x_5 \leq 72, \quad x_2 + 2x_3 + 4x_4 + 4x_5 \leq 17, \quad x_3 + x_4 + x_5 \leq 4.$$

The pack inequalities for $j = 4, 5$ are the upper bounds $x_4 \leq 1$ and $x_5 \leq 2$. \diamond

Under some additional assumptions on a and u along with the superincreasing property, one might be able to show that the packing inequality (7) is a strengthened integer cover or pack inequality of Atamtürk [21]. Our proof of Proposition 4 is direct, self-contained and motivated from the greedy solution.

When $u = \mathbf{e}$, we argue that (7) reduces to a minimal cover of the $0 \setminus 1$ superincreasing knapsack. Since $I = \{i \in N : \theta_i = 1\}$, inequality (7) becomes $x_j \leq 1$ for $j \in I$. For $j \notin I$, we have $\phi_j(i) = 1 \ \forall i \in I_j$ and (7) becomes $x_j + \sum_{i \in I_j} x_i \leq |I_j|$. The minimal covers can be obtained from [13, Theorem 2.4]: this theorem provides a set of integers $\kappa_1, \dots, \kappa_q$ for some $q \geq 1$ such that $\kappa_q = n$ and for any $i < q$, $\kappa_i := \max\{t < \kappa_{i+1} : \sum_{l=i+1}^n a_{\kappa_l} + a_t \leq b\}$. Proposition 2 gives us $\theta_l = 1$ if and only if $\sum_{k=l+1}^n a_k \theta_k < b$. Hence $\kappa_i = \max\{t < \kappa_{i+1} : 0 < a_t \leq b - \sum_{l=t+1}^n a_l \theta_l\} = \max\{t < \kappa_{i+1} : \theta_t = 1\}$. It follows that $\{\kappa_1, \dots, \kappa_q\} = I$. Theorem 2.5 in [13] states that any minimal cover is of the form $j \cup \{\kappa_i : \kappa_i > j\} = j \cup I_j$, for some $j \notin \{\kappa_1, \dots, \kappa_q\} = I$. Thus, the minimal cover inequalities are of the form $x_j + \sum_{i \in I_j} x_i \leq |I_j|, \ \forall j \notin I$, which is exactly the same as the packing inequalities.

The next result proves that the packing inequalities can be used to reformulate K . Later on in Theorem 2, we will prove that they also give an ideal formulation for K .

Proposition 5. $K = \left\{ x \in \mathcal{H}_n(u) : x_j + \sum_{i \in I_j} \phi_j(i)(x_i - \theta_i) \leq \theta_j, j = 1, \dots, n \right\}$.

Proof. The forward inclusion (\subseteq) is obvious due to the validity of inequalities (7). Consider $x \in \mathcal{H}_n(u)$ satisfying all the inequalities (7). Clearly, $x = \theta$ is a valid choice that belongs to K . Suppose that $x \neq \theta$ and define $i^* := \max\{i \in N : x_i \neq \theta_i\}$. Since x satisfies the inequality for $j = i^*$, $x_i = \theta_i \forall i > i^*$ implies $x_{i^*} \leq \theta_{i^*}$. Then $x_{i^*} \neq \theta_{i^*}$ gives us $x_{i^*} \leq \theta_{i^*} - 1$ and $x \preceq \theta$. Finally, Proposition 1 leads to $x \in K$. \square

We now show that the packing inequalities also define nontrivial facets of $\text{conv } K$. To aid our arguments, for every $j \in N$, we define $\xi_j : \mathbb{R}^n \mapsto \mathbb{R}$ as

$$\xi_j(x) := x_j - \theta_j + \sum_{i \in I_j} \phi_j(i)(x_i - \theta_i),$$

and the face defined by this inequality is $F_j := \{x \in \text{conv } K : \xi_j(x) = 0\}$. The integer points on this face have the following properties.

Proposition 6. Let $\hat{x} \in \mathcal{H}_n(u)$ be such that for some $j \in N$ with $\theta_j < u_j$ and $i \in I_j$, we have $x_k = \theta_k$ for all $k \in I_j$ with $k > i$.

- (1) If $\hat{x}_j = u_j$, $\hat{x}_k = u_k$ for all $k \in I_j$ with $k < i$ and $\hat{x}_i = \theta_i - 1$, then $\hat{x} \in F_j$.
- (2) If $\hat{x}_i \leq \theta_i - 2$, then $\hat{x} \notin F_j$.

Proof. For the first part, $\xi_j(\hat{x}) = u_j - \theta_j + \sum_{k \in I_j : k < i} \phi_j(k)(u_k - \theta_k) - \phi_j(i) = 0$, where the equality is due to Observation 1. Now suppose that $\hat{x}_i \leq \theta_i - 2$. Then

$$\xi_j(\hat{x}) \leq x_j - \theta_j + \sum_{k \in I_j : k < i} \phi_j(k)(x_k - \theta_k) - 2\phi_j(i) \leq u_j - \theta_j + \sum_{k \in I_j : k < i} \phi_j(k)(u_k - \theta_k) - 2\phi_j(i) = -\phi_j(i),$$

where the last equality is due to Observation 1. Now $\theta_j < u_j \implies \phi_j(i) > 0 \implies \xi_j(\hat{x}) < 0$. \square

Choosing $i = n$ in Proposition 6 yields the following inclusion that will be useful later in §4.1:

$$F_j \cap \mathbb{Z}^n \subseteq \{x \in \mathcal{H}_n(u) : x_n \in \{\theta_n - 1, \theta_n\}\} \quad \forall j \in N \text{ such that } \theta_j < u_j. \quad (8)$$

Proposition 7 (Facets). For any $j \in N$ with $\theta_j < u_j$, inequality (7) is facet-defining to $\text{conv } K$.

Proof. For $j \in N$ with $\theta_j < u_j$, we construct n affinely independent points of K that belong to $F_j \cap \mathbb{Z}^n$. These n points can be divided into three categories.

- (1) Fix $\hat{x} = (0, \dots, 0, \theta_j, \theta_{j+1}, \dots, \theta_n) \preceq \theta$. Clearly $\xi(\hat{x}) = 0$.
- (2) Fix $\hat{x} = (\mathbf{e}_l, u_j, 0, \dots, 0, \theta_{\text{next}(j)} - 1, \theta_{\text{next}(j)+1}, \dots, \theta_n) \preceq \theta$ for some $l < j$. Here $\xi(\hat{x}) = u_j - \theta_j - (u_j - \theta_j) = 0$.
- (3) Choose $i \in I_j$. There are two subtypes here: (a) fix $\hat{x} = (\mathbf{0}, u_j, \dots, u_{\text{prev}(i)}, \mathbf{0}, \theta_i - 1, \theta_{i+1}, \dots, \theta_n)$, (b) for some l such that $\max\{j, \text{prev}(i)\} < l < i$, fix $\hat{x} = (\mathbf{0}, u_j, \dots, u_{\text{prev}(i)}, \mathbf{e}_l, \theta_i - 1, \theta_{i+1}, \dots, \theta_n)$. Both these points satisfy $\hat{x} \preceq \theta$ by construction and are in F_j due to Proposition 6.

We have constructed a total of $1 + j - 1 + \text{next}(j) - j + \sum_{i \in I_j : i > \text{next}(j)} (i - \text{prev}(i)) = n$ points in F_j . Suppose that these n points form n columns of a matrix M in a way that the columns are sorted as Type 1, then Type 2, and then Type 3 (first all points of subtype (a) and then all of subtype (b)). Let there exist some weights $\lambda_1, \dots, \lambda_n$ such that $M\lambda = \mathbf{0}, \mathbf{e}^\top \lambda = 0$.

Claim 1. $\lambda_p = 0$ for all $p \notin \{j+1, \dots, j+|I_j|\}$. Consider the l^{th} row of M for some $1 \leq l < j$. There is exactly one column of M , corresponding to a Type 2 point, that contains a nonzero entry in row l . Hence $\lambda_2 = \dots = \lambda_j = 0$. Next consider some $k \notin I: k > j$. There is exactly one column of M , corresponding to a Type 3 subtype (b) point, that contains a nonzero entry in row k . Hence all the λ 's for Type 3 subtype (b) columns are zero, i.e. $\lambda_{j+|I_j|+1} = \dots = \lambda_n = 0$. The only remaining nonzero values are for $\lambda_1, \lambda_{j+1}, \dots, \lambda_{j+|I_j|}$, which must sum to zero. Consider the j^{th} row. Exactly one column of M has entry θ_j in row j (Type 1 point) while all other columns have entry u_j . This gives us $\lambda_1\theta_j + u_j \sum_{p=j+1}^{j+|I_j|} \lambda_p = 0$. Since $\lambda_1 + \sum_{p=j+1}^{j+|I_j|} \lambda_p = 0$ and $\theta_j < u_j$ by assumption, it follows that $\lambda_1 = \sum_{p=j+1}^{j+|I_j|} \lambda_p = 0$. \diamond

Claim 2. $\lambda = \mathbf{0}$. Consider the $\text{next}(j)^{th}$ row in M . The first Type 3 point has an entry $\theta_{\text{next}(j)} - 1$ in this row while all other Type 3 points have an entry of $u_{\text{next}(j)}$. The two equalities $\lambda_{j+1}(\theta_{\text{next}(j)} - 1) + \sum_{p=j+2}^{j+|I_j|} \lambda_p u_{\text{next}(j)} = 0$ and $\sum_{p=j+1}^{j+|I_j|} \lambda_p = 0$ imply $\lambda_{j+1}(\theta_{\text{next}(j)} - 1 - u_{\text{next}(j)}) = 0$, thereby giving us $\lambda_{j+1} = 0$ since $\theta_{\text{next}(j)} \leq u_{\text{next}(j)}$. Now let $i_t \in I_j \setminus \{\text{next}(j)\}$. Let the Type 3 subtype (a) point corresponding to i_t be in the $(j+t)^{th}$ column of M with the associated weight λ_{j+t} . Assume as part of induction hypothesis that $\lambda_{j+1} = \dots = \lambda_{j+t-1} = 0$. We argue that $\lambda_{j+t} = 0$. Observe that the entry for the i_t^{th} row of M in columns $j+t, j+t+1, \dots, j+|I_j|$ is $\theta_{i_t} - 1, u_{i_t}, \dots, u_{i_t}$, respectively. Upon using the induction hypothesis and $\lambda_{j+t} + \sum_{p=j+t+1}^{j+|I_j|} \lambda_p = 0$ in $(j+t)^{th}$ row of $M\lambda = \mathbf{0}$, we get $\lambda_{j+t}(\theta_{i_t} - 1 - u_{i_t}) = 0$, thereby giving us $\lambda_{j+t} = 0$ since $\theta_{i_t} \leq u_{i_t}$. This completes the induction process and we have $\lambda_p = 0, p = j+1, \dots, j+|I_j|$. Finally, $\lambda = \mathbf{0}$ follows from Claim 1. \diamond

We have shown in Claim 2 that $\lambda = \mathbf{0}$ is the only possible solution to $M\lambda = \mathbf{0}, \mathbf{e}^\top \lambda = 0$. Hence the n points constructed above are affinely independent and F_j is a facet of $\text{conv } K$. \square

Having shown that the packing inequalities (7) are strong valid inequalities for $\text{conv } K$, we now prove in Theorem 2 that $\text{conv } K$ does not have any other nontrivial facets. Our proof uses the dynamic program of Proposition 3 and Figure 1. For $j \in I$, recall \mathcal{L}_j , a subset of feasible solutions at the leaf child of j , from equation (4). We know from equation (5) that optimal solutions can only be found at leaves of the tree in Figure 1. If an optimal solution occurs at leaf \mathcal{L}_j , i.e. $\mathcal{L}_j \cap \Omega(c) \neq \emptyset$, we say that j is an optimal non-leaf node that is parent to the optimal leaf \mathcal{L}_j . While comparing two leaves \mathcal{L}_i and $\mathcal{L}_{i'}$, we say that \mathcal{L}_i is *larger* than $\mathcal{L}_{i'}$ if and only if $i > i'$.

Before proving Theorem 2, we present some useful characterizations of the optimal solutions of this dynamic program that will be invoked at multiple points in our proof. We will need the following notation: let the optimal value in Proposition 3 be stated as $f_j^*(c) = \max\{f_j^*(c)^-, f_j^*(c)^+\} \forall j \in I$, where

$$f_j^*(c)^- = [c_j]^+(\theta_j - 1) + \sum_{i=1}^{j-1} [c_i]^+ u_i, \quad f_j^*(c)^+ = c_j \theta_j + f_{\text{prev}(j)}^*(c).$$

It follows that $f_j^*(c)^- + \sum_{i \in I_j} c_i \theta_i = \max\{c^\top x : x \in \mathcal{L}_j\}$. Observe that since $x \leq u$, then $c_j = 0$ for some $j \in I$ implies that $f_j^*(c)^- \geq f_j^*(c)^+$. The next two observations are straightforward from the dynamic program of Figure 1.

Observation 2. For any $j \in I$, $\mathcal{L}_j \cap \Omega(c) \neq \emptyset \implies f_j^*(c)^- \geq f_j^*(c)^+$ and $f_j^*(c)^- > f_j^*(c)^+ \implies \Omega(c) \subseteq \mathcal{L}_j \cup \bigcup_{i \in I_j} \mathcal{L}_i$.

Observation 3. Let $j \in I$ and consider $x \in \mathcal{L}_j \cap \Omega(c)$. Then for any $i < j$, we have (i) $x_i = u_i$ if $c_i > 0$, (ii) $x_i = 0$ if $c_i < 0$, and (iii) x_i is unrestricted if $c_i = 0$.

The next result gives a sufficient condition for *all* the optimal solutions to lie in the facet defined by the j^{th} packing inequality.

Lemma 3. *Let $j \in N$ be such that $c_j > 0$, $c_t > 0 \forall t \in I_j$ and if $j \in I$, we also have $\mathcal{L}_j \cap \Omega(c) = \emptyset$. Then $\Omega(c) \subseteq F_j$.*

Proof. Recall that $\Omega(c) \subseteq \{\theta\} \cup \bigcup_{i \in I} \mathcal{L}_i$. We must show $\xi_j(x) = 0 \forall x \in \Omega(c)$. $\xi_j(\theta) = 0$ is trivial. Now consider $x \in \mathcal{L}_i \cap \Omega(c)$ for some $i \in I$. Our assumption $\mathcal{L}_j \cap \Omega(c) = \emptyset$ (in case $j \in I$) means that either $i \in I_j^-$ or $i \in I_j$. By construction of \mathcal{L}_i in (4), we have $x_t = \theta_t \forall t \in I_i$. If $i \in I_j^-$, then $x_j = \theta_j$ irrespective of whether $j \in I$ or not, and subsequently we have $\xi_j(x) = 0$. Now let $i \in I_j$. Since we assumed $c_j > 0$, $c_t > 0 \forall t \in I_j$, (4) gives us $x_i = \theta_i - 1$ and Observation 3 gives us $x_j = u_j$, $x_t = u_t \forall t \in I_j: t < i$. This along with $x_t = \theta_t \forall t \in I_i$ and Proposition 6 gives us $\xi_j(x) = 0$. \square

We are now ready to prove our first main result.

Theorem 2. $\text{conv } K = \left\{ x \in [\mathbf{0}, u]: x_j + \sum_{i \in I_j} \phi_j(i)(x_i - \theta_i) \leq \theta_j, \forall j \in N \right\}$.

Proof. Let $F_j = \{x \in \text{conv } K: \xi_j(x) = 0\}$, $\mathbb{U}_j = \{x \in \text{conv } K: x_j = u_j\}$, and $\mathbb{0}_j = \{x \in \text{conv } K: x_j = 0\}$ denote the faces of $\text{conv } K$ defined by the proposed inequalities. Note that $F_n = \mathbb{U}_n$ because $\theta_n = u_n$ and $I_n = \emptyset$. Based on Wolsey [22, Approach 6], we must show that for any $c \neq \mathbf{0}$, there exists $j \in N$ such that either $\Omega(c) \subseteq F_j$ or $\Omega(c) \subseteq \mathbb{U}_j$ or $\Omega(c) \subseteq \mathbb{0}_j$. If there exists $j \in N$ with $c_j < 0$, then clearly $\Omega(c) \subseteq \mathbb{0}_j$. Assume $c \geq \mathbf{0}$.

First suppose that $\mathcal{L}_n \cap \Omega(c) = \emptyset$. Let $i^* := \text{next}(\max\{j \in I: \mathcal{L}_j \cap \Omega(c) \neq \emptyset\})$ be the smallest non-leaf node that is larger than the parent of every optimal leaf node. Then for any $x \in \Omega(c)$, we have $x_j = \theta_j \forall j \in i^* \cup I_{i^*}$ and it follows that $\xi_j(x) = 0$ and hence $\Omega(c) \subseteq F_j$ for all $j \in i^* \cup I_{i^*}$. Henceforth assume $\mathcal{L}_n \cap \Omega(c) \neq \emptyset$. Recall that $\mathcal{L}_0 = \{\theta\}$.

Case i.: $\theta \notin \Omega(c)$. Let $i^* \in I$ be the parent node of the smallest optimal leaf. We first argue that there exists $i < i^*$ such that $c_i > 0$. Suppose $c_i = 0 \forall i < i^*$. Then $f_{i^*}^*(c)^- = c_{i^*}(\theta_{i^*} - 1) \leq c_{i^*}\theta_{i^*} = f_{i^*}^*(c)^+$ and hence $f_{i^*}^*(c) = f_{i^*}^*(c)^+ = c_{i^*}\theta_{i^*}$. Then every leaf \mathcal{L}_i , for all $i < i^*$ (including \mathcal{L}_0), is optimal, a contradiction to the optimality of \mathcal{L}_{i^*} . Hence there exists some $i < i^*$ such that $c_i > 0$. Since \mathcal{L}_{i^*} is the smallest optimal leaf and $i < i^*$, Observation 3 implies that $x_i = u_i \forall x \in \Omega(c)$ and thus $\Omega(c) \subseteq \mathbb{U}_i$.

Case ii.: $\theta \in \Omega(c)$.

Claim 3. $f_j^*(c) = f_j^*(c)^+ = c_j\theta_j + \sum_{i \in I_j^-} c_i\theta_i$ for all $j \in I$. Since $\theta \in \Omega(c)$, we have $f_{i_1}^*(c) = f_{i_1}^*(c)^+ = c_{i_1}\theta_{i_1}$. Let $j \in I \setminus \{i_1\}$. For any $i \in I_j^-$, Observation 2 and $\theta \in \Omega(c)$ give us $f_i^*(c)^- \leq f_i^*(c)^+$. This implies $f_i^*(c) = f_i^*(c)^+ = c_i\theta_i + f_{\text{prev}(i)}^*(c)$. Since we already argued $f_{i_1}^*(c) = c_{i_1}\theta_{i_1}$, a straightforward induction argument gives us the desired claim. \diamond

It follows that $f_j^*(c)^+ \geq f_j^*(c)^-$ for all $j \in I$.

Case ii-a.: $\exists j \notin I$ such that $c_j > 0$. We first argue that this case leads to $c_i > 0 \forall i \in I_j$. Suppose that $c_i = 0$ for some $i \in I_j$. Since $i > j$ with $j \notin I$, we have $j \notin I_i^-$. Consider the following:

$$f_i^*(c)^- - f_i^*(c)^+ = \sum_{t < i} c_t u_t - \sum_{t \in I_i^-} c_t \theta_t = \sum_{t \in I_i^-} c_t (u_t - \theta_t) + \sum_{t < i: t \notin I_i^- \cup j} c_t u_t + c_j u_j > 0,$$

where the strict inequality is due to $\theta \leq u, u > \mathbf{0}, c \geq \mathbf{0}$ and $c_j > 0$. Thus we have arrived at a contradiction to $f_i^*(c)^+ \geq f_i^*(c)^-$. Hence $c_i > 0 \forall i \in I_j$. Applying Lemma 3 gives us $\Omega(c) \subseteq F_j$.

Case ii-b.: $c_i = 0 \forall i \notin I$ or $I = \{1, \dots, n\}$. Since $c \neq \mathbf{0}$ and $c \geq \mathbf{0}$, there exists some $i \in I$ with $c_i > 0$. First suppose $c_i > 0$ for all $i \in I$. Since $f_{i_1}^*(c) = c_{i_1}\theta_{i_1}$ by Claim 3, we have $\mathcal{L}_{i_1} \cap \Omega(c) = \emptyset$

and Lemma 3 gives us $\Omega(c) \subseteq F_{i^*}$. Now let $i^* := \max\{i \in I : c_i = 0\}$ be the largest non-leaf node with cost coefficient equal to zero. By construction, $c_i > 0$ for all $i \in I_{i^*}$. The identity $f_{i^*}^*(c) = \max\{\sum_{i \in I_{i^*}^-} c_i u_i, \sum_{i \in I_{i^*}^-} c_i \theta_i\}$ along with $f_{i^*}^*(c) = \sum_{i \in I_{i^*}^-} c_i \theta_i$ from Claim 3 and $c_{i^*} = 0$ implies that

$$c_i = 0 \quad \text{OR} \quad (c_i > 0 \text{ and } \theta_i = u_i) \quad \forall i \in I_{i^*}^-. \quad (9)$$

First suppose there exists a $j \in I_{i^*}^-$ with $c_j > 0$. Equation (9) gives us $\theta_j = u_j$ and along with $I_j^- \subseteq I_{i^*}^-$, also implies $\sum_{t \in I_j^-} c_t (u_t - \theta_t) = 0$. Now $f_j^*(c)^+ = c_j + \sum_{t \in I_j^-} c_t \theta_t > \sum_{t \in I_j^-} c_t u_t = f_j^*(c)^-$ and as a result, Observation 2 implies $\mathcal{L}_j \cap \Omega(c) = \emptyset$. Consider an optimal solution $x \in \mathcal{L}_t \cap \Omega(c)$ for some $t \in \{0\} \cup (I \setminus \{j\})$. If $t > j$, then $c_j > 0$ and Observation 3 implies $x_j = u_j$. Otherwise $t < j$ and x_j is fixed to $\theta_j = u_j$. Hence $\Omega(c) \subseteq \cup_j$ if there exists a $j \in I_{i^*}^-$ with $c_j > 0$.

Finally, suppose that $c_i = 0 \forall i \in I_{i^*}^-$, or $i^* = i_1$ and $I_{i^*}^- = \emptyset$. Since we have already assumed in this case that $c_i = 0 \forall i \notin I$, it follows that $c_i = 0 \forall i < i^*$. Hence $i^* < n$, because otherwise $c = \mathbf{0}$. Consider $\text{next}(i^*)$, the first non-leaf node above i^* . The definition of i^* gives us $c_{\text{next}(i^*)} > 0$ and $c_i > 0 \forall i \in I_{\text{next}(i^*)}$. We argue that $\mathcal{L}_{\text{next}(i^*)} \cap \Omega(c) = \emptyset$; doing so and invoking Lemma 3 would lead to $\Omega(c) \subseteq F_{\text{next}(i^*)}$. Claim 3 gives us $f_{\text{next}(i^*)}^*(c) = f_{\text{next}(i^*)}^*(c)^+ = c_{\text{next}(i^*)} \theta_{\text{next}(i^*)}$ and since $c_{\text{next}(i^*)} > 0$, we have $f_{\text{next}(i^*)}^*(c)^+ > c_{\text{next}(i^*)} (\theta_{\text{next}(i^*)} - 1)$. Now $c_i = 0 \forall i < i^*$ implies that $f_{\text{next}(i^*)}^*(c)^- = c_{\text{next}(i^*)} (\theta_{\text{next}(i^*)} - 1)$. Hence $f_{\text{next}(i^*)}^*(c)^+ > f_{\text{next}(i^*)}^*(c)^-$ and Observation 2 implies that $\mathcal{L}_{\text{next}(i^*)} \cap \Omega(c) = \emptyset$.

All the above cases are mutually exclusive and exhaustive. Hence our proof is complete. \square

3.1. Applications of Theorem 2.

Lower bounded knapsack. For a lower bounded superincreasing knapsack $K^l := \{x \in [l, u] \cap \mathbb{Z}_+^n : a^\top x \leq b\}$, we can (i) perform a variable change $y = x - l$ to obtain $K = \{y \in \mathcal{H}_n(u - l) : a^\top y \leq b - a^\top l\}$, (ii) apply Theorem 2 to get $\text{conv } K$, and (iii) substitute back $x = y + l$ to obtain $\text{conv } K^l$. In particular, it is straightforward to verify that if $l \leq \theta$, then $\text{conv } K^l = (\text{conv } K) \cap [l, u]$.

Divisible knapsack. An integer basis is a strictly increasing sequence $\{a_i\}_{i \geq 1} \subset \mathbb{Z}_{++}$ with the property that there exists a sequence $\{u_i\}_{i \geq 1} \subset \mathbb{Z}_{++}$ such that every $b \in \mathbb{Z}_{++}$ can be expressed as $b = \sum_{i=1}^n a_i x_i$ for some n and $x \in \mathcal{H}_n(u)$. An equivalent characterization due to Cantor [cf. 23, Theorem 2.1] is the following: $\{a_i\}$ is an integer basis if and only if $a_1 = 1$ and $a_i \mid a_{i+1} \forall i$. Moreover, the sequence $\{u_i\}$ is uniquely determined as $u_i = \frac{a_{i+1}}{a_i} - 1$. Then, any finite subsequence of an integer basis $\{a_i\}$ and its corresponding $\{u_i\}$ define a divisible superincreasing knapsack, whose convex hull is given by Theorem 2. The set $K_\alpha(b)$ introduced in (1) is a particular case that uses powers of α as its integer basis. Another class of divisible superincreasing knapsacks is obtained when $u_i = u_0 \forall i$ and some $0 < u_0 \leq \frac{1}{2} \min_i \frac{a_{i+1}}{a_i}$.

4. INTERSECTION OF KNAPSACKS

In this section, we consider the problem of convexifying the intersection of $m \geq 2$ superincreasing knapsacks of \leq - or \geq -types. We prove that $\mathcal{O}(n)$ number of linear inequalities describe the convex hull of this intersection. The number of inequalities is independent of the number of intersecting knapsack sets. It suffices to address the case of *two* intersecting knapsacks; the general case follows immediately after noting that every superincreasing knapsack corresponds to a lexicographically ordered set of integer vectors and the lexicographic order is a total order. Our proof generalizes a recent result for $0 \setminus 1$ superincreasing knapsacks by Muldoon et al. [24].

Note that if we are given two \leq -type superincreasing knapsacks $\{x \in \mathcal{H}_n(u) : a^\top x \leq b\}$ with maximal packing γ and $\{x \in \mathcal{H}_n(u) : w^\top x \leq d\}$ with maximal packing θ , and w.o.l.o.g. we assume that $\gamma \preceq \theta$, then Proposition 1 tells us that their intersection is equal to $\{x \in \mathcal{H}_n(u) : x \preceq \gamma\}$. Hence the convex hull of the

intersection of two \leq -type knapsacks is given by $\mathcal{O}(n)$ packing inequalities corresponding to one of the sets. The nontrivial case to prove is when we are intersecting a \leq -type and a \geq -type knapsack.

Henceforth, let $K^{\leq} := \{x \in \mathcal{H}_n(u) : a^\top x \leq b\}$ and $K^{\geq} := \{x \in \mathcal{H}_n(u) : w^\top x \geq d\}$ be two superincreasing knapsacks with $a, w > \mathbf{0}$ and $b, d > 0$. Proposition 1 implies that

$$K^{\leq} \cap K^{\geq} = \{x \in \mathcal{H}_n(u) : \gamma \preceq x \preceq \theta\}, \quad (10)$$

where θ is the maximal packing of K given by (2) and γ is the minimal packing of K^{\geq} obtained from (2) by complementing variables:

$$\gamma_i := u_i - \min \left\{ u_i, \left\lfloor \frac{w^\top u - d - \sum_{k=i+1}^n w_k(u_k - \gamma_k)}{w_i} \right\rfloor \right\} \quad \forall i = n, \dots, 1.$$

It follows from Theorem 2 that

$$\text{conv } K^{\geq} = \left\{ x \in [\mathbf{0}, u] : x_j + \sum_{i \in T_j} \Phi_j(i)(x_i - \gamma_i) \geq \gamma_j \quad \forall j \in N \right\}, \quad (11)$$

where $T = \{i \in N : \gamma_i \leq u_i - 1\}$, $T_j = \{i \in T : i > j\}$, and $\Phi_j(i) = \gamma_j \prod_{k \in T_j : k < i} (\gamma_k + 1)$ for all $i \in T_j$. The main result of this section proves that the convex hull operator distributes over $K^{\leq} \cap K^{\geq}$.

Theorem 3. $\text{conv } \{x \in \mathcal{H}_n(u) : \gamma \preceq x \preceq \theta\} = \text{conv } \{x \in \mathcal{H}_n(u) : x \succcurlyeq \gamma\} \cap \text{conv } \{x \in \mathcal{H}_n(u) : x \preceq \theta\}$. In particular, if $a, w > \mathbf{0}$, then $\text{conv } (K^{\leq} \cap K^{\geq}) = \text{conv } K^{\leq} \cap \text{conv } K^{\geq}$.

This is an interesting result because in general for any two arbitrary sets \mathcal{X}_1 and \mathcal{X}_2 , we have $\text{conv } (\mathcal{X}_1 \cap \mathcal{X}_2) \subseteq \text{conv } \mathcal{X}_1 \cap \text{conv } \mathcal{X}_2$. For the intersection $\mathcal{X} = \bigcap_{i=1}^{m_1} K_i^{\leq} \cap \bigcap_{i=1}^{m_2} K_i^{\geq}$ of $m \geq 2$ superincreasing knapsacks (each having a coefficient vector of strictly positive integers), the equivalence to lexicographic ordering implies that $\mathcal{X} = K_s^{\leq} \cap K_{s'}^{\geq}$ for some indices s, s' . Then Theorem 3 gives us $\mathcal{O}(n)$ inequalities to describe the convex hull of this intersection.

Remark 3. For $K^{\leq} \cap K^{\geq}$, the assumption $a, w > \mathbf{0}$ is not w.o.l.o.g since we are considering two knapsacks simultaneously. Suppose that $a_i > 0$ for all $i \in N_1 := \{1, \dots, n_1\}$ and some $1 \leq n_1 \leq n-1$, and $w_i > 0$ for all $i \in N_2 \subseteq N$ (assume w.o.l.o.g. that $\{n_1 + 1, \dots, n\} \subseteq N_2$). In this case, we don't have the distributive property and in general, $\text{conv } (K^{\leq} \cap K^{\geq}) \subsetneq \text{conv } K^{\leq} \cap \text{conv } K^{\geq}$, as shown by the following example.

Example 2. Let $K^{\leq} = \{x \in \mathbb{Z}_+^7 : 2x_1 + 8x_2 + 46x_3 + 150x_4 + 310x_5 \leq 841, x \leq (3, 5, 2, 1, 2, 4, 2)\}$ and $K^{\geq} = \{x \in \mathbb{Z}_+^7 : 2x_4 + 7x_5 + 30x_6 + 50x_7 \geq 150, x \leq (3, 5, 2, 1, 2, 4, 2)\}$. The inequalities describing $\text{conv } K^{\leq}$ are described in Example 1 in §3 whereas the nontrivial facets of $\text{conv } K^{\geq}$ are $x_5 + 2x_6 + 4x_7 \geq 12$, $x_6 + x_7 \geq 4$ and $x_7 \geq 1$. The PORTA software [25] tells us that the intersection of $\text{conv } K^{\leq}$ and $\text{conv } K^{\geq}$ has a fractional extreme point $(0, 0, 2, 1, 1, 7/2, 1)$. \diamond

In fact, we claim that $K^{\leq} \cap K^{\geq}$ may not be equal to the set of integer points that are lexicographically ordered between two given integer vectors. Suppose it were true: $K^{\leq} \cap K^{\geq} = \{x \in \mathcal{H}_n(u) : \gamma' \preceq x \preceq \theta'\}$ for some $\gamma', \theta' \in \mathbb{Z}_+^n$. Then, because $K^{\leq} = \{x \in \mathcal{H}_n(u) : (x_i)_{i \in N_1} \preceq (\theta_i)_{i \in N_1}\}$ and $K^{\geq} = \{x \in \mathcal{H}_n(u) : (x_i)_{i \in N_2} \succcurlyeq (\gamma_i)_{i \in N_2}\}$, we must have $\theta'_i = \theta_i \quad \forall i \in N_1, \theta'_i = u_i \quad \forall i \in N \setminus N_1, \gamma'_i = \gamma_i \quad \forall i \in N_2, \gamma'_i = 0 \quad \forall i \in N \setminus N_2$. It is obvious that $K^{\leq} \cap K^{\geq} \subseteq \{x \in \mathcal{H}_n(u) : \gamma' \preceq x \preceq \theta'\}$. Since the knapsack $\sum_{i \in N_1} a_i x_i \leq b$ is nontrivial (i.e. $\sum_{i \in N_1} a_i u_i > b$), there must exist some $k \in \{1, \dots, n_1 - 1\}$ such that $\theta_k < u_k$. Now, $x' := (\gamma'_1, \dots, \gamma'_{k-1}, \max\{\theta_k + 1, \gamma'_k\}, \gamma'_{k+1}, \dots, \gamma'_{n-1}, \theta_n - 1)$ satisfies $\gamma' \preceq x' \preceq \theta'$ but $(x'_i)_{i \in N_1} \not\preceq (\theta_i)_{i \in N_1}$. In fact, $x' \notin \text{conv } K^{\leq}$. This gives us a contradiction. Thus, in the presence of zeros in the coefficients of at least one of the two knapsacks, we cannot use the nice structural properties of lexicographic orderings to convexify the intersection of K^{\leq} and K^{\geq} .

The rest of this section is devoted to proving Theorem 3. We assume throughout that $a, w > \mathbf{0}$ and hence identity (10) holds true. In order to ensure that $K^{\leq} \cap K^{\geq}$ is a full-dimensional set, we assume w.o.l.o.g. that $\sum_{i=1}^{n-1} w_i u_i + (\theta_n - 1)w_n \geq d$ and $\gamma_n \leq \theta_n - 1$; otherwise we can fix $x_n = \theta_n$ and address the lower dimensional case with modified right hand sides. Our arguments are divided into two cases: $\gamma_n \leq \theta_n - 2$ and $\gamma_n = \theta_n - 1$. The proof of the first case depends on a geometric intuition, as explained in §4.1. This geometric insight breaks down when $\gamma_n = \theta_n - 1$ and hence we resort to some technical lemmas in §4.2. Based on these building blocks, the proof of Theorem 3 is presented in §4.3.

4.1. $\gamma_n \leq \theta_n - 2$. Consider Figure 2. It is apparent that in the two-dimensional case, we always have $\text{conv}\{x \in \mathcal{H}_2(u): \gamma \preceq x \preceq \theta\} = \text{conv}\{x \in \mathcal{H}_2(u): x \succcurlyeq \gamma\} \cap \{x \in \mathcal{H}_2(u): x \preceq \theta\}$. In Figure 2(a), where $\gamma_2 \leq \theta_2 - 2$, we see that $\text{conv}\{x \in \mathcal{H}_2(u): \gamma \preceq x \preceq \theta\}$ is equal to $(A \cup B) \cup (B \cup C)$, where the three sets are defined as follows: $A = \{x \in \text{conv}\{y \in \mathcal{H}_2(u): y \preceq \theta\}, x_2 \geq \theta_2 - 1\}$, $B = [0, u_1] \times [\gamma_2 + 1, \theta_2 - 1]$ and $C = \{x \in \text{conv}\{y \in \mathcal{H}_2(u): y \succcurlyeq \gamma\}, x_2 \leq \gamma_2 + 1\}$. This geometric intuition of expressing the convex hull as a union of two sets enables us to prove that the convex hull operator distributes for arbitrary n when $\gamma_n \leq \theta_n - 2$.

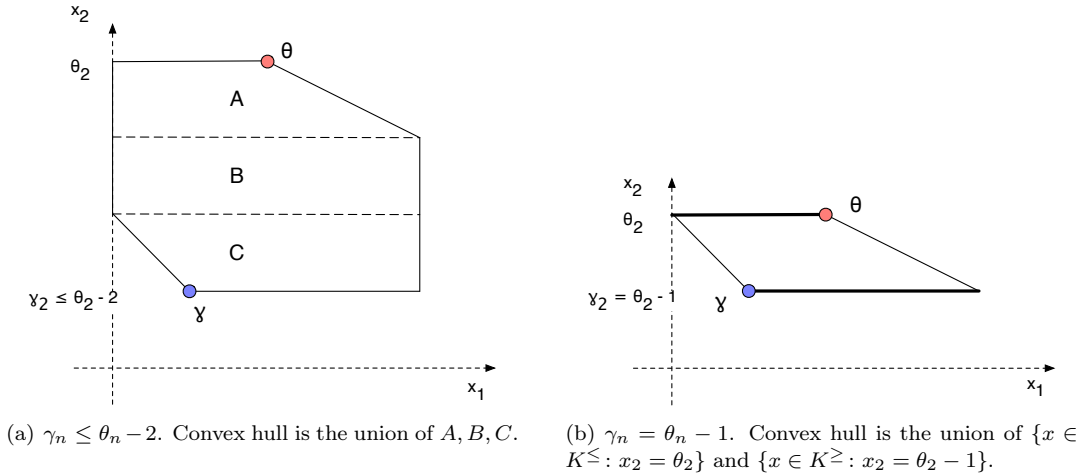


FIGURE 2. Two cases for convexifying $\{x \in \mathcal{H}_n(u): \gamma \preceq x \preceq \theta\}$.

Proposition 8. Assume that $\gamma_n \leq \theta_n - 2$. Then

$$\text{conv}(K^{\leq} \cap K^{\geq}) = \{x \in \text{conv } K^{\leq}: x_n \geq \gamma_n + 1\} \cup \{x \in \text{conv } K^{\geq}: x_n \leq \theta_n - 1\} = \text{conv } K^{\leq} \cap \text{conv } K^{\geq}.$$

Proof. $\text{conv}(K^{\leq} \cap K^{\geq}) \subseteq \text{conv } K^{\leq} \cap \text{conv } K^{\geq}$ is obvious. Take $y \in \text{conv } K^{\leq} \cap \text{conv } K^{\geq}$. If $y_n \geq \gamma_n + 1$ then $y \in \text{conv } K^{\leq} \cap \{x: x_n \geq \gamma_n + 1\}$; otherwise $y_n < \gamma_n + 1$ and the assumption $\gamma_n + 1 \leq \theta_n - 1$ implies that $y \in \text{conv } K^{\geq} \cap \{x: x_n \leq \theta_n - 1\}$. Now let y be an extreme point of $\text{conv } K^{\leq} \cap \{x: x_n \geq \gamma_n + 1\}$. Theorem 2 and equation (8) imply that $y \in K^{\leq}$ with $y_n \in \{\gamma_n + 1, \theta_n - 1, \theta_n\}$. Since $\gamma_n + 1 \leq \theta_n - 1$, it follows that $\gamma \preceq y \preceq \theta$ and identity (10) establishes $y \in \text{conv}(K \cap K^{\geq})$. The arguments for $y \in \text{conv } K^{\geq} \cap \{x: x_n \leq \theta_n - 1\}$ are similar. Thus $\{x \in \text{conv } K^{\leq}: x_n \geq \gamma_n + 1\} \cup \{x \in \text{conv } K^{\geq}: x_n \leq \theta_n - 1\} \subseteq \text{conv}(K^{\leq} \cap K^{\geq})$, thereby completing our proof. \square

The above proof heavily relies on the assumption $\gamma_n \leq \theta_n - 2$. In particular, if $\gamma_n = \theta_n - 1$, then we can only show that $\{x \in \text{conv } K^{\leq}: x_n \geq \gamma_n + 1\} \cup \{x \in \text{conv } K^{\geq}: x_n \leq \theta_n - 1\} \subseteq \text{conv } K^{\leq} \cap \text{conv } K^{\geq}$ but

cannot argue the \supseteq -inclusion using the above steps. Hence the case $\gamma_n = \theta_n - 1$ requires a different proof technique, which is presented next.

4.2. $\gamma_n = \theta_n - 1$. We start by writing a disjunctive representation of $K^\leq \cap K^\geq$, somewhat similar in vein to the first equality in Proposition 8.

Lemma 4. $K^\leq \cap K^\geq = \{x \in K^\leq : x_n = \theta_n\} \cup \{x \in K^\geq : x_n = \theta_n - 1\}$.

Proof. Since $x_n \leq u_n = \theta_n$ for any $x \in K^\leq \cap K^\geq$, the elementary disjunction $\{x_n \leq \theta_n - 1\} \cup \{x_n = \theta_n\}$ gives us $K^\leq \cap K^\geq = \{x \in K^\leq \cap K^\geq : x_n \leq \theta_n - 1\} \cup \{x \in K^\leq \cap K^\geq : x_n = \theta_n\}$. For any $x \in \mathcal{H}_n(u)$ such that $x_n \leq \theta_n - 1$, we have $x \prec \theta$ and hence $K^\leq \cap \{x : x_n \leq \theta_n - 1\} = \mathcal{H}_n(u) \cap \{x : x_n \leq \theta_n - 1\}$. Since $K^\geq = \{x \in \mathcal{H}_n(u) : x \succ \gamma\}$ and $\gamma_n = \theta_n - 1$, we get $K^\leq \cap K^\geq \cap \{x : x_n \leq \theta_n - 1\} = K^\geq \cap \{x : x_n = \theta_n - 1\}$. Next, note that $w_n \theta_n \geq \sum_{i=1}^{n-1} w_i u_i + w_n(\theta_n - 1) \geq d$. Then it follows that $\{x \in \mathcal{H}_n(u) : x_n = \theta_n\} \subset K^\geq$ and we get $K^\leq \cap K^\geq \cap \{x : x_n = \theta_n\} = K^\leq \cap \{x : x_n = \theta_n\}$. \square

Lemma 4 will be crucial in completing the proof of this case in §4.3. The proposed disjunction is depicted in Figure 2(b) for \mathbb{R}^2 . It is easy to see that the two nontrivial facets obtained by convexifying this union in \mathbb{R}^2 are exactly the packing inequalities for $\{x \in \mathcal{H}_2(u) : x \preceq \theta\}$ and $\{x \in \mathcal{H}_2(u) : x \succ \gamma\}$. Motivated by this illustration, our approach is use the extended formulation of Balas [26] to convexify the union in Lemma 4 and argue that every fractional point in $\text{conv } K^\leq \cap \text{conv } K^\geq$, i.e. $x \in \text{conv } K^\leq \cap \text{conv } K^\geq$ such that $x_n = \theta_n - \epsilon$ for some $\epsilon \in (0, 1)$, belongs to the convex hull of $K^\leq \cap K^\geq$. To do so, we must characterize points in the ϵ -restrictions of $\text{conv } K^\leq$ and $\text{conv } K^\geq$. This is achieved in the next lemma. Recall from Definition 2 that we denote the support of θ as $I = \{i_1, \dots, i_r, i_{r+1} = n\}$ for some $r \geq 0$.

Lemma 5. Let $\epsilon \in (0, 1)$ and $x \in [\mathbf{0}, u]$ such that $x_n = \theta_n - \epsilon$. Define $\delta_i := \min\{\epsilon u_i, x_i\}$ for all $i \leq i_r$ and $\delta_i := x_i$ for all $i > i_r$.

- (1) If $x \in \text{conv } K^\leq$, then $x_j - \delta_j - \theta_j(1 - \epsilon) + \sum_{i \in I_j \setminus n} \phi_j(i)(x_i - \delta_i - \theta_i(1 - \epsilon)) \leq 0$ for every $j \in N \setminus n$.
- (2) If $x \in \text{conv } K^\geq$, then $\delta_j + \sum_{i \in T_j \setminus n} \Phi_j(i)\delta_i \geq \epsilon \Phi_j(n)$ for every $j \in N \setminus n$.

To prove this technical lemma, we need to exploit the recursive nature of $\phi_j(\cdot)$ and $\Phi_j(\cdot)$ so that we can rearrange expressions suitably. The following lemma gives us the required result; its proof is a tedious algebraic exercise and is hence relegated to A.

Lemma 6. Let $j \in N \setminus n$ and $\epsilon \neq 0$.

- (1) For $s, i \in I_j$ with $s < i$, $\phi_j(i) = \phi_j(s) \left[1 + \sum_{s \leq k < i} \phi_k(i) \right]$.
- (2) For $s \in I_j \setminus n$, $\sum_{i \in I_j \setminus n, i \geq s} \phi_j(i)(x_i - z_i - \theta_i \epsilon) = \phi_j(s) \sum_{i \in I_j \setminus n, i \geq s} \left[x_i - z_i - \theta_i \epsilon + \sum_{k \in I_i \setminus n} \phi_i(k)(x_k - z_k - \theta_k \epsilon) \right]$.
- (3) For $s \in T_j \setminus n$, $\sum_{i \in T_j \setminus n, i \geq s} \Phi_j(i)(\gamma_i \epsilon - z_i) = \Phi_j(s) \sum_{i \in T_j \setminus n, i \geq s} \left[\epsilon \Phi_i(n) - z_i - \sum_{k \in T_i \setminus n} \Phi_i(k)z_k \right]$.

Proof. In A. \square

Proof of Lemma 5. We prove the first part here; arguments for the second part are analogous and are provided in A for completeness. Choose $j \in N \setminus n$. If $\theta_j = u_j$, then the inequality is obvious because $\phi_j(\cdot) = 0$. Assume $\theta_j < u_j$. For $j \geq i_r$, the inequality holds because $I_j \setminus n = \emptyset$ and the j^{th} nontrivial facet in $\text{conv } K^\leq \cap \{x : x_n = \theta_n - \epsilon\}$ can be written as $x_j - \epsilon u_j \leq \theta_j(1 - \epsilon)$. Now consider $j \leq i_r - 1$ and assume that the inequality holds for all $i = j + 1, \dots, n - 1$.

Claim 4. For any $s \in I_j \setminus n$, we have $\sum_{i \in I_j \setminus n : i \geq s} \phi_j(i)(x_i - z_i - \theta_i(1 - \epsilon)) \leq 0$. Follows from Lemma 6, $\phi_j(s) \geq 0$ and induction hypothesis. \diamond

First suppose that $z_j = \epsilon u_j$. Let $s \in I_j \setminus n$ be the smallest index such that $z_s = x_s$. If s does not exist, then $z_i = \epsilon u_i \forall i \in I_j \setminus n$ and upon rearranging, we must show that

$$x_j + \sum_{i \in I_j \setminus n} \phi_j(i)(x_i - \theta_i) - \epsilon \left[u_j - \theta_j + \sum_{i \in I_j \setminus n} \phi_j(i)(u_i - \theta_i) \right] \leq \theta_j,$$

which is exactly the j^{th} nontrivial facet of $\text{conv } K^{\leq} \cap \{x: x_n = \theta_n - \epsilon\}$ due to $\phi_j(n) = u_j - \theta_j + \sum_{i \in I_j \setminus n} \phi_j(i)(u_i - \theta_i)$ from Observation 1. Now suppose s exists. Rewriting the desired inequality, we must show that

$$\frac{x_j - \theta_j}{u_j - \theta_j} - \epsilon \leq \Lambda := - \sum_{i \in I_j \setminus n} \frac{\phi_j(i)}{u_j - \theta_j} (x_i - z_i - \theta_i(1 - \epsilon)).$$

It suffices to verify that $\Lambda \geq 1 - \epsilon$. We split the summation over $I_j \setminus n$ into three parts: first over all $i < s$, then s , and third over all $i > s$, and use the first part of Lemma 6 to rewrite $\phi_j(i)/(u_j - \theta_j) \forall i \in I_j \setminus n$. This gives us

$$\begin{aligned} \Lambda = & - \sum_{i \in I_j: i < s} \left[1 + \sum_{k \in I_j: k < i} \phi_k(i) \right] (x_i - \epsilon u_i - \theta_i(1 - \epsilon)) + \left[1 + \sum_{k \in I_j: k < s} \phi_k(s) \right] \theta_s(1 - \epsilon) \\ & - \sum_{i \in I_j \setminus n: i > s} \left[1 + \sum_{k \in I_j: k < s} \phi_k(i) + \sum_{\substack{k \in I_j: \\ s \leq k < i}} \phi_k(i) \right] (x_i - z_i - \theta_i(1 - \epsilon)). \end{aligned}$$

Combining common terms gives us

$$\begin{aligned} \Lambda = & \theta_s(1 - \epsilon) - \Lambda_1 - \Lambda_2 := \theta_s(1 - \epsilon) - \sum_{i \in I_j: i < s} \left[x_i - \epsilon u_i - \theta_i(1 - \epsilon) + \sum_{t \in I_i \setminus n} \phi_i(t)(x_t - z_t - \theta_t(1 - \epsilon)) \right] \\ & - \sum_{i \in I_j \setminus n: i > s} \left[1 + \sum_{\substack{k \in I_j: \\ s \leq k < i}} \phi_k(i) \right] (x_i - z_i - \theta_i(1 - \epsilon)). \end{aligned}$$

The induction hypothesis implies $\Lambda_1 \leq 0$. First statement of Lemma 6 implies $\Lambda_2 = \frac{1}{\phi_j(s)} \sum_{i \in I_j \setminus n: i > s} \phi_j(i)(x_i - z_i - \theta_i(1 - \epsilon))$. Claim 4 now gives us $\Lambda_2 \leq 0$. Thus $\Lambda \geq \theta_s(1 - \epsilon) \geq 1 - \epsilon$ since $s \in I_j$ and $\theta_s \geq 1$.

Finally, let $z_j = x_j$. Then Claim 4 gives us $-\theta_j(1 - \epsilon) + \sum_{i \in I_j \setminus n} \phi_j(i)(x_i - y_i - \theta_i(1 - \epsilon)) \leq 0$. \square

4.3. Proving the distributive property.

Proof of Theorem 3. The case $\gamma_n \leq \theta_n - 2$ is proved in Proposition 8. Now suppose that $\gamma_n = \theta_n - 1$. It remains to show that $\text{conv}(K^{\leq} \cap K^{\geq}) \supseteq \text{conv } K^{\leq} \cap \text{conv } K^{\geq}$ since the \subseteq -inclusion is obvious. We first obtain an extended formulation for $\text{conv}(K^{\leq} \cap K^{\geq})$. For convenience, define

$$g_j := \theta_j + \sum_{i \in I_j \setminus n} \phi_j(i)\theta_i \quad \forall j \in N: \theta_j \leq u_j - 1, \quad h_j := \gamma_j + \sum_{i \in T_j \setminus n} \Phi_j(i)\gamma_i \quad \forall j \in N: \gamma_j \geq 1. \quad (12)$$

Claim 5. $\text{conv}(K^{\leq} \cap K^{\geq})$ is equal to the projection onto the x -space of the polytope

$$\begin{aligned} \mathcal{P} := & \left\{ (x, y) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : \mathbf{0} \leq y \leq x, \theta_n - 1 \leq x_n \leq \theta_n, y_n = (\theta_n - 1)(\theta_n - x_n), x_i = y_i, i_r < i < n \right. \\ & y_i + \frac{u_i}{\theta_n}(x_n - y_n) \leq u_i, \quad i < n, \quad y_i - x_i + \frac{u_i}{\theta_n}(x_n - y_n) \geq 0, \quad i \leq i_r \quad (13a) \\ & x_j - y_j + \sum_{i \in I_j \setminus n} \phi_j(i)(x_i - y_i) - \frac{g_j}{\theta_n}(x_n - y_n) \leq 0, \quad j \in N : \theta_j < u_j \quad (13b) \\ & \left. y_j + \sum_{i \in T_j \setminus n} \Phi_j(i)y_i + \frac{h_j}{\theta_n}(x_n - y_n) \geq h_j, \quad j \in N : \gamma_j \geq 1 \right\}. \quad (13c) \end{aligned}$$

The proof of this claim makes use of Lemma 4 and the disjunctive programming result of Balas [26]; it is provided in A. \diamond

Since $\gamma_n = \theta_n - 1$, we get $\text{conv}\{x \in K^{\geq} : x_n = \theta_n - 1\} = \text{conv} K^{\geq} \cap \{x : x_n = \gamma_n\}$. The fact that $x_n \leq \theta_n$ defines a face of $\text{conv} K^{\leq}$ implies that $\text{conv}\{x \in K^{\leq} : x_n = \theta_n\} = \text{conv} K^{\leq} \cap \{x : x_n = \theta_n\}$. Consider $x \in \text{conv} K^{\leq} \cap \text{conv} K^{\geq}$. Note that since $x_n \leq \theta_n$ is valid to $\text{conv} K^{\leq}$ and $x_n \geq \gamma_n = \theta_n - 1$ is valid to $\text{conv} K^{\geq}$, it must be that $x_n \in [\theta_n - 1, \theta_n]$. If $x_n \in \{\theta_n - 1, \theta_n\}$, then Lemma 4 implies $x \in \text{conv}(K^{\leq} \cap K^{\geq})$. Let $x_n = \theta_n - \epsilon$ for some $\epsilon \in (0, 1)$. Fix $y \in \mathbb{R}_+^n$ as follows: $y_i = \min\{\epsilon u_i, x_i\}$ for $i \leq i_r$, $y_i = x_i$ for $i_r < i < n$, and $y_n = \epsilon(\theta_n - 1)$. From Claim 5, it suffices to show that $(x, y) \in \mathcal{P}$. By construction, y satisfies (13a) and the trivial relations with x . Since $x_n - y_n = \theta_n(1 - \epsilon)$, (13b) and (13c), respectively, are transformed to

$$x_j - y_j + \sum_{i \in I_j \setminus n} \phi_j(i)(x_i - y_i) \leq g_j(1 - \epsilon) \quad \forall j \in N : \theta_j \leq u_j - 1 \quad (14a)$$

$$y_j + \sum_{i \in T_j \setminus n} \Phi_j(i)y_i \geq h_j\epsilon \quad \forall j \in N : \gamma_j \geq 1. \quad (14b)$$

Since $h_j = \Phi_j(n)$ (analogous to Observation 1), inequality (14b) becomes $y_j + \sum_{i \in T_j \setminus n} \Phi_j(i)y_i \geq \epsilon\Phi_j(n)$. Then Lemma 5 (with y replacing z) implies that (14a) and (14b) are satisfied. Hence $(x, y) \in \mathcal{P}$. \square

Remark 4. We believe that the proof used for the difficult case $\gamma_n = \theta_n - 1$ can be modified to handle the case $\gamma_n \leq \theta_n - 2$ as well. However our geometric arguments in §4.1 lend more intuition into the structural properties of intersection of \succ and \preccurlyeq ordered cones.

4.4. Application of Theorem 3. Consider a mixed integer knapsack with a single continuous variable defined by the set $Q := \{(x, y) \in \mathcal{H}_n(u) \times [0, \tilde{u}] : \sum_{i \in N} a_i x_i + y \leq b\}$, where we assume that $\{(a_i, u_i)\}_{i \in N}$ forms a superincreasing sequence of tuples of positive integers and \tilde{u} and b are positive reals with $\tilde{u} \leq b$. It is straightforward to verify that $Q = Q_1 \cup Q_2$, where

$$\begin{aligned} Q_1 &:= \{(x, y) \in \mathcal{H}_n(u) \times \mathbb{R} : \lceil b - \tilde{u} \rceil \leq a^\top x \leq \lfloor b \rfloor, 0 \leq y \leq b - a^\top x\} \\ Q_2 &:= \{(x, y) \in \mathcal{H}_n(u) \times [0, \tilde{u}] : a^\top x \leq \lfloor b - \tilde{u} \rfloor\}. \end{aligned}$$

Let $\theta_{\lfloor b \rfloor}$ and $\theta_{\lfloor b - \tilde{u} \rfloor}$ denote maximal packings for $a^\top x \leq \lfloor b \rfloor$ and $a^\top x \leq \lfloor b - \tilde{u} \rfloor$, respectively, and $\gamma_{\lceil b - \tilde{u} \rceil}$ denote a minimal packing for $a^\top x \geq \lceil b - \tilde{u} \rceil$. It follows that $\text{conv} Q_2 = \text{conv}\{x \in \mathcal{H}_n(u) : x \preccurlyeq \theta_{\lfloor b - \tilde{u} \rfloor}\} \times [0, \tilde{u}]$ and hence Theorem 2 gives us

$$\text{conv} Q_2 = \{(x, y) \in [\mathbf{0}, u] \times [0, \tilde{u}] : \text{inequalities (7) for } \theta_{\lfloor b - \tilde{u} \rfloor}\}. \quad (15)$$

For the convex hull of Q_1 , observe that $b - a^\top x \geq 0$ is valid to $\text{conv } Q_1$ and hence

$$\begin{aligned} \text{conv } Q_1 &= (\text{conv} \{(x, y) \in \mathcal{H}_n(u) \times \mathbb{R} : \gamma_{\lceil b - \bar{u} \rceil} \preceq x \preceq \theta_{\lfloor b \rfloor}\}) \cap \{(x, y) : 0 \leq y \leq b - a^\top x\} \\ &= \{(x, y) \in [0, u] \times \mathbb{R}_+ : \text{inequalities (7) for } \theta_{\lfloor b \rfloor}, \text{ inequalities (11) for } \gamma_{\lceil b - \bar{u} \rceil}, a^\top x + y \leq b\} \end{aligned} \quad (16)$$

where the second equality is from Theorem 3. Since $Q = Q_1 \cup Q_2$, we have $\text{conv } Q = \text{conv}(\text{conv } Q_1 \cup \text{conv } Q_2)$. Equations (15) and (16) and disjunctive programming [26] imply a compact extended formulation for $\text{conv } Q$.

5. DISCUSSION

In this paper, we have identified a special class of general integer knapsacks, referred to as superincreasing knapsacks. We studied its greedy solution θ and showed that this well-structured set is equal to the set of integer vectors that are lexicographically less (\preceq) than θ . The convex hull of this \preceq -ordered set is described using $\mathcal{O}(n)$ facets, where n is the dimension of the knapsack, and all the nontrivial facets are derived from θ . An arbitrary knapsack is in general a strict subset of solutions that are \preceq -than the greedy solution and hence our facet description yields a class of valid inequalities that can be possibly strengthened by other means for use in cutting plane algorithms. A second interesting phenomenon exhibited by the \preceq -ordering and superincreasing structure is that the convex hull operator distributes over a finite intersection. Our results generalize previously known descriptions for 0/1 superincreasing knapsacks.

Generalized lexicographic cone. Finally, we mention that the results derived in this paper can be generalized as follows. Given $\beta \in \mathbb{Z}_{++}^n$, let $x \preceq_\beta \theta$ denote that x is β -lex smaller than θ , i.e. either $x = \theta$ or the first index i in reverse order is such that $x_i \leq \theta_i - \beta_i$. Suppose that we are interested in convexifying $\mathcal{C} := \{x \in \mathcal{H}_n(u) : x \preceq_\beta \theta\}$. It is easy to verify that if $\beta \neq \mathbf{e}$, then \mathcal{C} cannot be represented as a integer knapsack; we may need a disjunctive formulation to include the correct set of feasible solutions. We believe that by exploiting the properties of lexicographic orderings, all the results proved in this paper carry through with suitable adjustments; for example the function $\phi_j(\cdot)$ in equation (6) must be modified to $\phi_j(i) := \frac{(u_j - \theta_j)}{\beta_j} \prod_{k=\text{next}(j)}^{\text{prev}(i)} : \frac{(u_k + \beta_k - \theta_k)}{\beta_k}$ for all $i \in I_j$.

REFERENCES

- [1] L. Wolsey, Faces for a linear inequality in 0–1 variables, *Mathematical Programming* 8 (1975) 165–178.
- [2] M. Padberg, (1, k)-configurations and facets for packing problems, *Mathematical Programming* 18 (1980) 94–99.
- [3] R. Weismantel, On the 0/1 knapsack polytope, *Mathematical Programming* 77 (1997) 49–68.
- [4] R. Weismantel, Hilbert bases and the facets of special knapsack polytopes, *Mathematics of Operations Research* 21 (1996) 886–904.
- [5] O. Marcotte, The cutting stock problem and integer rounding, *Mathematical Programming* 33 (1985) 82–92.
- [6] Y. Pochet, L. Wolsey, Integer knapsack and flow covers with divisible coefficients: polyhedra, optimization and separation, *Discrete Applied Mathematics* 59 (1995) 57–74.
- [7] Y. Pochet, R. Weismantel, The sequential knapsack polytope, *SIAM Journal on Optimization* 8 (1998) 248–264.
- [8] V. Cacchiani, A. Caprara, G. Maróti, P. Toth, On integer polytopes with few nonzero vertices, *Operations Research Letters* 41 (2013) 74 – 77.
- [9] R. Merkle, M. Hellman, Hiding information and signatures in trapdoor knapsacks, *IEEE Transactions on Information Theory* 24 (1978) 525–530.
- [10] A. M. Odlyzko, The rise and fall of knapsack cryptosystems, *Cryptology and Computational Number Theory* 42 (1990) 75–88.

- [11] A. Shamir, A polynomial-time algorithm for breaking the basic Merkle-Hellman cryptosystem, *IEEE Transactions on Information Theory* 30 (1984) 699–704.
- [12] M. Atkinson, A. Negro, N. Santoro, Sums of lexicographically ordered sets, *Discrete Mathematics* 80 (1990) 115–122.
- [13] M. Laurent, A. Sassano, A characterization of knapsacks with the max-flow–min-cut property, *Operations Research Letters* 11 (1992) 105–110.
- [14] A. Gupte, S. Ahmed, M. Cheon, S. Dey, Solving mixed integer bilinear problems using MILP formulations, *SIAM Journal on Optimization* 23 (2013) 721–744.
- [15] R. Gillmann, V. Kaibel, Revlex-initial 0/1-polytopes, *Journal of Combinatorial Theory, Series A* 113 (2006) 799–821.
- [16] A. Martin, R. Weismantel, The intersection of knapsack polyhedra and extensions, in: R. E. Bixby, E. A. Boyd, R. Z. Rios-Mercado (Eds.), *Integer Programming and Combinatorial Optimization*, volume 1412 of *Lecture Notes in Computer Science*, Springer, 1998, pp. 243–256.
- [17] Q. Louveaux, R. Weismantel, Polyhedral properties for the intersection of two knapsacks, *Mathematical Programming* 113 (2008) 15–37.
- [18] E. Fernández, K. Jørnsten, Partial cover and complete cover inequalities, *Operations Research Letters* 15 (1994) 19–33.
- [19] M. Magazine, G. L. Nemhauser, L. E. Trotter, When the greedy solution solves a class of knapsack problems, *Operations Research* 23 (1975) 207–217.
- [20] J. L. R. Alfonsin, On variations of the subset sum problem, *Discrete Applied Mathematics* 81 (1998) 1–7.
- [21] A. Atamtürk, Cover and pack inequalities for (mixed) integer programming, *Annals of Operations Research* 139 (2005) 21–38.
- [22] L. Wolsey, Strong valid inequalities, in: *Integer Programming, Discrete Mathematics and Optimization*, Wiley-Interscience, 1998, pp. 145–147.
- [23] P. Pitteloud, On the log-concavity of sequences arising from integer bases, *Discrete Applied Mathematics* 138 (2004) 153–175.
- [24] F. Muldoon, W. Adams, H. Sherali, Ideal representations of lexicographic orderings and base-2 expansions of integer variables, *Operations Research Letters* 41 (2013) 32–39.
- [25] T. Christof, A. Löbel, PORTA: POLYhedron Representation Transformation Algorithm, Available electronically from http://typo.zib.de/opt-long_projects/Software/Porta/, 1997. Lastchecked: November 2012.
- [26] E. Balas, Disjunctive programming: Properties of the convex hull of feasible points, *Discrete Applied Mathematics* 89 (1998) 3–44.

APPENDIX A. MISSING PROOFS OF §4

Proof of Lemma 6. For the first part, note that $\frac{\phi_j(i)}{\phi_j(s)} = \prod_{s \leq k < i} (u_k + 1 - \theta_k) = 1 + \sum_{s \leq k < i} (u_k - \theta_k) \prod_{k < t < i} (u_t + 1 - \theta_t) = 1 + \sum_{s \leq k < i} \phi_k(i)$. We prove the second statement by induction on $|\{i \in I_j \setminus n : i \geq s\}|$. The third statement can be proven similarly via induction on $|\{i \in T_j \setminus n : i \geq s\}|$. The claim is clearly true when the cardinality is 1. Assume it is true when the cardinality is $m \geq 1$ and let $|\{i \in I_j \setminus n : i \geq s\}| = m + 1$. For convenience, denote $\{i \in I_j \setminus n : i \geq s\} = \{s_1 := s, s_2, \dots, s_{m+1}\}$. Then

the left hand side in the lemma is

$$\begin{aligned} & \sum_{i \in \{s_1, \dots, s_m\}} \phi_j(i)(x_i - z_i - \theta_i \epsilon) + \phi_j(s_{m+1})(x_{s_{m+1}} - z_{s_{m+1}} - \theta_{s_{m+1}} \epsilon) \\ &= \phi_j(s) \sum_{i \in \{s_1, \dots, s_m\}} \left[x_i - z_i - \theta_i \epsilon + \sum_{k \in I_i \setminus \{n, s_{m+1}\}} \phi_i(k)(x_k - z_k - \theta_k \epsilon) \right] + \phi_j(s_{m+1})(x_{s_{m+1}} - z_{s_{m+1}} - \theta_{s_{m+1}} \epsilon) \end{aligned}$$

where the equality is obtained by invoking the induction hypothesis on the first term. Substituting $\phi_j(s_{m+1}) = \phi_j(s) [1 + \sum_{t=1}^m \phi_{s_t}(s_{m+1})]$ from the first part into the above equality and combining common terms, we get

$$\sum_{\substack{i \in I_j \setminus n \\ i \geq s}} \phi_j(i)(x_i - z_i - \theta_i \epsilon) = \phi_j(s) \sum_{i \in \{s_1, \dots, s_{m+1}\}} \left[x_i - z_i - \theta_i \epsilon + \sum_{k \in I_i \setminus n} \phi_i(k)(x_k - z_k - \theta_k \epsilon) \right],$$

which is the desired result. \square

Proof of second part of Lemma 5. If $\gamma_j = 0$, then $\Phi_j(\cdot) = 0$ and again the inequality is obvious. For $j > i_r$, we have $z_i = x_i \forall i \in j \cup T_j \setminus n$. Then $x \in \text{conv } K^{\geq} \cap \{x: x_n = \theta_n - \epsilon\}$ implies $x_j + \sum_{i \in T_j \setminus n} \Phi_j(i)x_i \geq \gamma_j + \sum_{i \in T_j} \Phi_j(i)\gamma_i - \Phi_j(n)(\theta_n - \epsilon)$ and the right hand side can be simplified to $\epsilon\Phi_j(n)$ (analogous to Observation 1) and $\gamma_n = \theta_n - 1$. Now consider $j \leq i_r$ and assume that the inequality holds for all $i = j+1, \dots, n-1$. Let $s \in T_j \setminus n$ be the smallest index such that $z_s = \epsilon u_s$ and suppose that s exists. Upon rearranging terms, we have to check that

$$\min\{\epsilon u_j, x_j\} + \sum_{i \in T_j: i < s} \Phi_j(i)x_i \geq \Lambda := \epsilon\Phi_j(n) - \Phi_j(s)\epsilon u_s - \sum_{i \in T_j: s < i < n} \Phi_j(i)z_i.$$

It suffices to verify that $\Lambda \leq 0$.

Claim 6. For any $j \in N \setminus n$, we have $\Phi_j(n) = \gamma_j + \sum_{i \in T_j \setminus n} \Phi_j(i)\gamma_i = \Phi_j(k)(\gamma_k + 1) + \sum_{i \in T_k \setminus n} \Phi_j(i)\gamma_i$ for all $k \in T_j \setminus n$. The first equality is analogous to Observation 1. The second statement follows from a straightforward reverse induction on k and using the fact that $\Phi_j(\text{next}(k)) = \Phi_j(k)(\gamma_k + 1)$. \diamond

Using Claim 6 with $k = s$, we rewrite $\Phi_j(n)$ to get $\Lambda = \epsilon\Phi_j(s)(\gamma_s + 1 - u_s) + \sum_{i \in T_j: s < i < n} \Phi_j(i)(\epsilon\gamma_i - z_i)$. Applying Lemma 6 with $\epsilon = \epsilon$ gives us

$$\Lambda = \epsilon\Phi_j(s)(\gamma_s + 1 - u_s) + \Phi_j(\text{next}(s)) \sum_{\substack{i \in T_j \\ s < i < n}} \left[\epsilon\Phi_i(n) - z_i - \sum_{k \in T_i \setminus n} \Phi_i(k)z_k \right] \leq \epsilon\Phi_j(s)(\gamma_s + 1 - u_s) \leq 0$$

where $\text{next}(s) = \min\{i: i \in T_j\}$, the first inequality is due to each summand being non-positive from induction hypothesis and the second inequality is due to $s \in T_j$ and hence $\gamma_s \leq u_s - 1$.

If s does not exist then $y_i = x_i \forall i \in T_j \setminus n$ and we must show that $\min\{\epsilon u_j, x_j\} + \sum_{i \in T_j \setminus n} \Phi_j(i)x_i \geq \epsilon\Phi_j(n)$. If $x_j \leq \epsilon u_j$, then the same argument as that used for $j > i_r$ proves the desired inequality. Otherwise $\epsilon u_j < x_j$. In this case, we set $\Lambda = \epsilon\Phi_j(n) - \sum_{i \in T_j \setminus n} \Phi_j(i)x_i$, rewrite $\Phi_j(n) = \gamma_j + \sum_{i \in T_j \setminus n} \Phi_j(i)\gamma_i$ as in Claim 6 and follow same steps as before to obtain $\Lambda \leq \epsilon\gamma_j \leq \epsilon u_j$, as desired. \square

Proof of Claim 5. We have $\text{conv}(K^{\leq} \cap K^{\geq}) = \text{conv}(\text{conv}\{x \in K^{\leq}: x_n = \theta_n\} \cup \text{conv}\{x \in K^{\geq}: x_n = \theta_n - 1\})$ from Lemma 4. Since $\gamma_n = \theta_n - 1$, we get $\text{conv}\{x \in K^{\geq}: x_n = \theta_n - 1\} = \text{conv } K^{\geq} \cap \{x: x_n = \gamma_n\}$. The fact that $x_n \leq \theta_n$ defines a face of $\text{conv } K^{\leq}$ implies that $\text{conv}\{x \in K^{\leq}: x_n = \theta_n\} = \text{conv } K^{\leq} \cap \{x: x_n = \theta_n\}$.

Applying Balas' result and invoking Theorem 2 and equation (11) gives us the following extended formulation for $\text{conv}(K^{\leq} \cap K^{\geq})$:

$$\begin{aligned} \mathcal{P}' = \Big\{ (x, y, \lambda) \in \mathbb{R}_+^n \times \mathbb{R}_+^n \times [0, 1] : & \quad y_n = \gamma_n(1 - \lambda), \quad \mathbf{0} \leq y \leq u(1 - \lambda), \quad \mathbf{0} \leq x - y \leq u\lambda, \\ & \quad x_n - y_n = \theta_n \lambda \quad x_i - y_i = 0 \quad i_r < i < n, \\ & \quad x_j - y_j + \sum_{i \in I_j \setminus n} \phi_j(i)(x_i - y_i) \leq g_j \lambda, \quad j \leq i_r : \theta_j \leq u_j - 1 \\ & \quad y_j + \sum_{i \in T_j \setminus n} \Phi_j(i)y_i \geq h_j(1 - \lambda), \quad j \in N : \gamma_j \geq 1 \Big\}. \end{aligned}$$

where g_j and h_j are defined in (12). The equality $x_n - y_n = \theta_n \lambda$ implies $\lambda = (x_n - y_n)/\theta_n$. Upon substituting for λ in \mathcal{P}' and rearranging the inequalities, we get the proposed claim. \square